

DOCKET NO.: 4028

USPS PRESS MAIL
EL 547 896 345 US
NOVEMBER 20 2000

INVENTORS: Christoph STAHL
Thomas FECHNER
Oliver ROCKINGER

TITLE OF THE INVENTION

METHOD FOR RECOGNIZING OBJECTS IN AN IMAGE PIXEL PLANE

PRIORITY CLAIM

This application is based on and claims the priority under 35 U.S.C. §119 of German Patent Application 199 55 919.8, filed on November 20, 1999, the entire disclosure of which is incorporated herein by reference.

FIELD OF THE INVENTION

The invention relates to a method for recognizing in an input image objects belonging to at least one given object class. The images are provided as digital images. The recognition takes place in or on the image pixel plane. The present method is particularly suitable for the automatic evaluation of image data representing large quantities of image informations.

BACKGROUND INFORMATION

Large volumes of image data are available in many fields of use. Such large image data volumes must be analyzed in accordance with predetermined criteria. For example, in the area of military reconnaissance, it is frequently the case that large quantities of image data of scenes or terrains are acquired by sensors. These acquired image data must be scrutinized with regard to the presence of installations, vehicles, infrastructure features and so forth in the terrain. These image data are generally acquired in large numbers which must be processed and evaluated within given time limitations. The objects to be recognized may have any random dimensions and may have a structure that characterizes any particular object. The structure of the object may be rather complex or it may be simple. In all these cases it is desirable to perform an automatic image analysis as rapidly as possible.

Other fields of application of this type of image evaluation are, for example, to be found in the area of medical diagnosis, for example when it is necessary to examine a large number of X-ray images, for instance for recognizing anomalies such as tumors or the like. Another example where an automatic image analysis method is employed is in the area of police work. This area includes the search for missing persons, the monitoring of border crossings or the like. In all these areas a reliable automatic rapid image analysis method provides great advantages.

General, theoretical approaches for such analysis method for the recognition of objects in images are known from an article in "Technical Report ISIS TR-4" by T. Dodd, University of South Hampton, 1996. This article describes different possible approaches to the analysis of digital images for the purpose of recognizing objects in such images.

10

00000000000000000000000000000000

15

20

25

Individual steps for analyzing images are known from the following publications. Different methods for a rough classification of objects are described in an article "Classifier and Shift-Invariant Automatic Target Recognition Neural Networks", by D. P. Casasent, L. M. Neiberg published in "Neural Networks", Vol. 8, No. 7/8, by Elsevier Science Ltd., 1995. General methods for the dissecting or decomposing of a digital image into image components represented by signals are found, for example in a publication "Practice of Digital Image Processing and Pattern Recognition" by P. Haberaecker, published by Carl Hanser Verlag, 1995. The so-called "Ensemble Theory for Classifiers" has been described in an article "Machine Learning Research" by T. G. Dietterich that appeared in "AI Magazine", Vol. 18, No. 4, 1997, published by AAAI Press. A possible way of merging or fusing individual results of an analysis are described in "Vehicle Detection in Infrared Line Scan Imagery Using Belief Networks" by P. G. Dubksbury, D. M. Booth and C. J. Radford, published at the 5th International Conference of Image Processing and Application, Edinburgh, 1995.

4028/WGF:ar

German Patent Publication DE 44 38 235 A1 discloses a method for recognizing objects in natural surroundings. The known method uses several classifiers which operate in accordance with a predetermined, simple rule. A disadvantage of the known method is seen in that it can function only, and on principle, under the assumptions that the objects to be recognized are compact and at least partially symmetric. Thus, the method is useful only for the recognition of point-shaped objects. The method cannot be used for recognizing larger and structurized objects.

German Patent Publication DE 196 39 884 C2 discloses a system for recognizing patterns in images. For classifying an object, the known system requires the input and processing of feature or characterizing parameters of the object in order to produce a recognition result based on such parameters. The system requires the use of a complex classifier. A disadvantage of the known system is seen in that only a clearly determined object can be recognized.

OBJECTS OF THE INVENTION

In view of the above it is the aim of the invention to achieve the following objects singly or in combination:

to provide an image analyzing method that operates sufficiently reliable for the intended purposes of simultaneously

recognizing a multitude of complex objects of any random size in an image;

INS
AI to provide an image recognition method or process which, by its very systematic, is suitable for an automatic 5 object recognition from a large number of images under time limitations;

to provide a recognition method that yields as the result of its analyzing procedure an output statement regarding the presence of a multitude and complex object in an image; and

to provide a rapid and automatic evaluation of large quantities of image data by a concrete and robust analysis method for an object recognition.

SUMMARY OF THE INVENTION

According to the invention at least one object class or a plurality of different object classes and respective classifiers are used simultaneously. Thus, it is possible to evaluate an input image for a respective number of object classes. In a fusing or merging step of reduced images a decision is made for each object class and that decision provides directly the 15 position and type of the recognized object or objects. In order to achieve such a rapid analyzation of a multitude of images the 20 invention employs rules that are learned by a neural network on

the basis of representative examples. A characterizing or feature vector is used for the formation of the classifiers. The feature vector is formed from a previously determined vicinity of a respective relevant pixel point of a corresponding filter image.

More specifically, the method according to the invention comprises the following steps:

- Mild
P-5-03
- 10
15
20
25
- (a) roughly classifying (10) pixel points of said received images whether or not a pixel point is relevant for said object recognition to provide relevant pixel points;
 - (b) forming (11) a reduced image based on relevant pixel points as roughly classified in step (a);
 - (c) filtering (20) each reduced image (11) for forming at least two corresponding decomposed or filtered images (21, 22, 23) whereby image components relevant for said object recognition are retained in said filtered images;
 - (d) further classifying (30) said filtered images for providing classified images,, wherein said further classifying is performed by a group of different classifiers which operate in accordance with learned rules to allocate said classified images to different object classes, wherein each of said classifiers operates based on a characterizing vector forming an input information for its respective classifier;
 - (e) merging or fusing (40) said classified images in accordance with an algorithm to form a combined global evaluation or

decision for each class of said object classes, said global evaluation or decision representing merged images (41A, 41B, 41C);

- 5 (f) deciding (50), on the basis of said merged images, whether a pixel point is relevant and if so to which of said object classes each relevant pixel point belongs.

10 The method according to the invention is robust and practical and can be used for the recognition of a large number of complex, different object classes, such as vehicles, persons, buildings and so forth. Furthermore, the present method is suitable for use in connection with a large number of different image sources such as video cameras, infrared cameras, X-ray cameras and so forth.

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
165
170
175
180
185
190
195
200
205
210
215
220
225
230
235
240
245
250
255
260
265
270
275
280
285
290
295
300
305
310
315
320
325
330
335
340
345
350
355
360
365
370
375
380
385
390
395
400
405
410
415
420
425
430
435
440
445
450
455
460
465
470
475
480
485
490
495
500
505
510
515
520
525
530
535
540
545
550
555
560
565
570
575
580
585
590
595
600
605
610
615
620
625
630
635
640
645
650
655
660
665
670
675
680
685
690
695
700
705
710
715
720
725
730
735
740
745
750
755
760
765
770
775
780
785
790
795
800
805
810
815
820
825
830
835
840
845
850
855
860
865
870
875
880
885
890
895
900
905
910
915
920
925
930
935
940
945
950
955
960
965
970
975
980
985
990
995
1000
1005
1010
1015
1020
1025
1030
1035
1040
1045
1050
1055
1060
1065
1070
1075
1080
1085
1090
1095
1100
1105
1110
1115
1120
1125
1130
1135
1140
1145
1150
1155
1160
1165
1170
1175
1180
1185
1190
1195
1200
1205
1210
1215
1220
1225
1230
1235
1240
1245
1250
1255
1260
1265
1270
1275
1280
1285
1290
1295
1300
1305
1310
1315
1320
1325
1330
1335
1340
1345
1350
1355
1360
1365
1370
1375
1380
1385
1390
1395
1400
1405
1410
1415
1420
1425
1430
1435
1440
1445
1450
1455
1460
1465
1470
1475
1480
1485
1490
1495
1500
1505
1510
1515
1520
1525
1530
1535
1540
1545
1550
1555
1560
1565
1570
1575
1580
1585
1590
1595
1600
1605
1610
1615
1620
1625
1630
1635
1640
1645
1650
1655
1660
1665
1670
1675
1680
1685
1690
1695
1700
1705
1710
1715
1720
1725
1730
1735
1740
1745
1750
1755
1760
1765
1770
1775
1780
1785
1790
1795
1800
1805
1810
1815
1820
1825
1830
1835
1840
1845
1850
1855
1860
1865
1870
1875
1880
1885
1890
1895
1900
1905
1910
1915
1920
1925
1930
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
2045
2050
2055
2060
2065
2070
2075
2080
2085
2090
2095
2100
2105
2110
2115
2120
2125
2130
2135
2140
2145
2150
2155
2160
2165
2170
2175
2180
2185
2190
2195
2200
2205
2210
2215
2220
2225
2230
2235
2240
2245
2250
2255
2260
2265
2270
2275
2280
2285
2290
2295
2300
2305
2310
2315
2320
2325
2330
2335
2340
2345
2350
2355
2360
2365
2370
2375
2380
2385
2390
2395
2400
2405
2410
2415
2420
2425
2430
2435
2440
2445
2450
2455
2460
2465
2470
2475
2480
2485
2490
2495
2500
2505
2510
2515
2520
2525
2530
2535
2540
2545
2550
2555
2560
2565
2570
2575
2580
2585
2590
2595
2600
2605
2610
2615
2620
2625
2630
2635
2640
2645
2650
2655
2660
2665
2670
2675
2680
2685
2690
2695
2700
2705
2710
2715
2720
2725
2730
2735
2740
2745
2750
2755
2760
2765
2770
2775
2780
2785
2790
2795
2800
2805
2810
2815
2820
2825
2830
2835
2840
2845
2850
2855
2860
2865
2870
2875
2880
2885
2890
2895
2900
2905
2910
2915
2920
2925
2930
2935
2940
2945
2950
2955
2960
2965
2970
2975
2980
2985
2990
2995
3000
3005
3010
3015
3020
3025
3030
3035
3040
3045
3050
3055
3060
3065
3070
3075
3080
3085
3090
3095
3100
3105
3110
3115
3120
3125
3130
3135
3140
3145
3150
3155
3160
3165
3170
3175
3180
3185
3190
3195
3200
3205
3210
3215
3220
3225
3230
3235
3240
3245
3250
3255
3260
3265
3270
3275
3280
3285
3290
3295
3300
3305
3310
3315
3320
3325
3330
3335
3340
3345
3350
3355
3360
3365
3370
3375
3380
3385
3390
3395
3400
3405
3410
3415
3420
3425
3430
3435
3440
3445
3450
3455
3460
3465
3470
3475
3480
3485
3490
3495
3500
3505
3510
3515
3520
3525
3530
3535
3540
3545
3550
3555
3560
3565
3570
3575
3580
3585
3590
3595
3600
3605
3610
3615
3620
3625
3630
3635
3640
3645
3650
3655
3660
3665
3670
3675
3680
3685
3690
3695
3700
3705
3710
3715
3720
3725
3730
3735
3740
3745
3750
3755
3760
3765
3770
3775
3780
3785
3790
3795
3800
3805
3810
3815
3820
3825
3830
3835
3840
3845
3850
3855
3860
3865
3870
3875
3880
3885
3890
3895
3900
3905
3910
3915
3920
3925
3930
3935
3940
3945
3950
3955
3960
3965
3970
3975
3980
3985
3990
3995
4000
4005
4010
4015
4020
4025
4030
4035
4040
4045
4050
4055
4060
4065
4070
4075
4080
4085
4090
4095
4100
4105
4110
4115
4120
4125
4130
4135
4140
4145
4150
4155
4160
4165
4170
4175
4180
4185
4190
4195
4200
4205
4210
4215
4220
4225
4230
4235
4240
4245
4250
4255
4260
4265
4270
4275
4280
4285
4290
4295
4300
4305
4310
4315
4320
4325
4330
4335
4340
4345
4350
4355
4360
4365
4370
4375
4380
4385
4390
4395
4400
4405
4410
4415
4420
4425
4430
4435
4440
4445
4450
4455
4460
4465
4470
4475
4480
4485
4490
4495
4500
4505
4510
4515
4520
4525
4530
4535
4540
4545
4550
4555
4560
4565
4570
4575
4580
4585
4590
4595
4600
4605
4610
4615
4620
4625
4630
4635
4640
4645
4650
4655
4660
4665
4670
4675
4680
4685
4690
4695
4700
4705
4710
4715
4720
4725
4730
4735
4740
4745
4750
4755
4760
4765
4770
4775
4780
4785
4790
4795
4800
4805
4810
4815
4820
4825
4830
4835
4840
4845
4850
4855
4860
4865
4870
4875
4880
4885
4890
4895
4900
4905
4910
4915
4920
4925
4930
4935
4940
4945
4950
4955
4960
4965
4970
4975
4980
4985
4990
4995
5000
5005
5010
5015
5020
5025
5030
5035
5040
5045
5050
5055
5060
5065
5070
5075
5080
5085
5090
5095
5100
5105
5110
5115
5120
5125
5130
5135
5140
5145
5150
5155
5160
5165
5170
5175
5180
5185
5190
5195
5200
5205
5210
5215
5220
5225
5230
5235
5240
5245
5250
5255
5260
5265
5270
5275
5280
5285
5290
5295
5300
5305
5310
5315
5320
5325
5330
5335
5340
5345
5350
5355
5360
5365
5370
5375
5380
5385
5390
5395
5400
5405
5410
5415
5420
5425
5430
5435
5440
5445
5450
5455
5460
5465
5470
5475
5480
5485
5490
5495
5500
5505
5510
5515
5520
5525
5530
5535
5540
5545
5550
5555
5560
5565
5570
5575
5580
5585
5590
5595
5600
5605
5610
5615
5620
5625
5630
5635
5640
5645
5650
5655
5660
5665
5670
5675
5680
5685
5690
5695
5700
5705
5710
5715
5720
5725
5730
5735
5740
5745
5750
5755
5760
5765
5770
5775
5780
5785
5790
5795
5800
5805
5810
5815
5820
5825
5830
5835
5840
5845
5850
5855
5860
5865
5870
5875
5880
5885
5890
5895
5900
5905
5910
5915
5920
5925
5930
5935
5940
5945
5950
5955
5960
5965
5970
5975
5980
5985
5990
5995
6000
6005
6010
6015
6020
6025
6030
6035
6040
6045
6050
6055
6060
6065
6070
6075
6080
6085
6090
6095
6100
6105
6110
6115
6120
6125
6130
6135
6140
6145
6150
6155
6160
6165
6170
6175
6180
6185
6190
6195
6200
6205
6210
6215
6220
6225
6230
6235
6240
6245
6250
6255
6260
6265
6270
6275
6280
6285
6290
6295
6300
6305
6310
6315
6320
6325
6330
6335
6340
6345
6350
6355
6360
6365
6370
6375
6380
6385
6390
6395
6400
6405
6410
6415
6420
6425
6430
6435
6440
6445
6450
6455
6460
6465
6470
6475
6480
6485
6490
6495
6500
6505
6510
6515
6520
6525
6530
6535
6540
6545
6550
6555
6560
6565
6570
6575
6580
6585
6590
6595
6600
6605
6610
6615
6620
6625
6630
6635
6640
6645
6650
6655
6660
6665
6670
6675
6680
6685
6690
6695
6700
6705
6710
6715
6720
6725
6730
6735
6740
6745
6750
6755
6760
6765
6770
6775
6780
6785
6790
6795
6800
6805
6810
6815
6820
6825
6830
6835
6840
6845
6850
6855
6860
6865
6870
6875
6880
6885
6890
6895
6900
6905
6910
6915
6920
6925
6930
6935
6940
6945
6950
6955
6960
6965
6970
6975
6980
6985
6990
6995
7000
7005
7010
7015
7020
7025
7030
7035
7040
7045
7050
7055
7060
7065
7070
7075
7080
7085
7090
7095
7100
7105
7110
7115
7120
7125
7130
7135
7140
7145
7150
7155
7160
7165
7170
7175
7180
7185
7190
7195
7200
7205
7210
7215
7220
7225
7230
7235
7240
7245
7250
7255
7260
7265
7270
7275
7280
7285
7290
7295
7300
7305
7310
7315
7320
7325
7330
7335
7340
7345
7350
7355
7360
7365
7370
7375
7380
7385
7390
7395
7400
7405
7410
7415
7420
7425
7430
7435
7440
7445
7450
7455
7460
7465
7470
7475
7480
7485
7490
7495
7500
7505
7510
7515
7520
7525
7530
7535
7540
7545
7550
7555
7560
7565
7570
7575
7580
7585
7590
7595
7600
7605
7610
7615
7620
7625
7630
7635
7640
7645
7650
7655
7660
7665
7670
7675
7680
7685
7690
7695
7700
7705
7710
7715
7720
7725
7730
7735
7740
7745
7750
7755
7760
7765
7770
7775
7780
7785
7790
7795
7800
7805
7810
7815
7820
7825
7830
7835
7840
7845
7850
7855
7860
7865
7870
7875
7880
7885
7890
7895
7900
7905
7910
7915
7920
7925
7930
7935
7940
7945
7950
7955
7960
7965
7970
7975
7980
7985
7990
7995
8000
8005
8010
8015
8020
8025
8030
8035
8040
8045
8050
8055
8060
8065
8070
8075
8080
8085
8090
8095
8100
8105
8110
8115
8120
8125
8130
8135
8140
8145
8150
8155
8160
8165
8170
8175
8180
8185
8190
8195
8200
8205
8210
8215
8220
8225
8230
8235
8240
8245
8250
8255
8260
8265
8270
8275
8280
8285
8290
8295
8300
8305
8310
8315
8320
8325
8330
8335
8340
8345
8350
8355
8360
8365
8370
8375
8380
8385
8390
8395
8400
8405
8410
8415
8420
8425
8430
8435
8440
8445
8450
8455
8460
8465
8470
8475
8480
8485
8490
8495
8500
8505
8510
8515
8520
8525
8530
8535
8540
8545
8550
8555
8560
8565
8570
8575
8580
8585
8590
8595
8600
8605
8610
8615
8620
8625
8630
8635
8640
8645
8650
8655
8660
8665
8670
8675
8680
8685
8690
8695
8700
8705
8710
8715
8720
8725
8730
8735
8740
8745
8750
8755
8760
8765
8770
8775
8780
8785
8790
8795
8800
8805
8810
8815
8820
8825
8830
8835
8840
8845
8850
8855
8860
8865
8870
8875
8880
8885
8890
8895
8900
8905
8910
8915
8920
8925
8930
8935
8940
8945
8950
8955
8960
8965
8970
8975
8980
8985
8990
8995
9000
9005
9010
9015
9020
9025
9030
9035
9040
9045
9050
9055
9060
9065
9070
9075
9080
9085
9090
9095
9100
9105
9110
9115
9120
9125
9130
9135
9140
9145
9150
9155
9160
9165
9170
9175
9180
9185
9190
9195
9200
9205
9210
9215
9220
9225
9230
9235
9240
9245
9250
9255
9260
9265
9270
9275
9280
9285
9290
9295
9300
9305
9310
9315
9320
9325
9330
9335
9340
9345
9350
9355
9360
9365
9370
9375
9380
9385
9390
9395
9400
9405
9410
9415
9420
9425
9430
9435
9440
9445
9450
9455

Fig. 2 shows in its left side an example of a received or input image that has been reduced in a first step of the present method to form a reduced image shown on the right side;

5 Fig. 3 shows symbolically a possible decision result image achieved by the last step of the present method showing pixel points that are not relevant by a zero and recognition relevant pixel clusters designated by a respective classification number; and

10 Fig. 4 is a flow diagram similar to that of Fig. 1, however illustrating the present method as applied to a single object class.

15 DETAILED DESCRIPTION OF PREFERRED EXAMPLE EMBODIMENTS AND OF THE BEST MODE OF THE INVENTION

20 Referring to Fig. 1, the method according to the invention is performed for recognizing objects in images. Three object classes are processed in the flow diagram of Fig. 1. First, an input or individual image 1 shown on the left-hand side of Fig. 2 is separately analyzed in a first rough classification step 10. In this rough classification step 10 the input image 1 is viewed as a unit. Stated differently, the input image 1 is not yet analyzed or dissected. However, in step 10 the input image 1 is

separated by into pixel points that are relevant for an object recognition and pixel points that are irrelevant for the object recognition and thus can be ignored.

The right side of Fig. 2 shows the rough classification. Dark areas show relevant pixels, light areas show irrelevant pixels.

A suitably selected rough classifier determines the relevance or irrelevance of each pixel of the input image 1 for the following method steps. The individual image points of the input image 1 are roughly classified in accordance with a given first criterion that determines the weightiness or significance of the individual pixel point for the object classification and recognition. This given first criterion takes into account, for example whether the image has interesting areas, which may be recognizable objects or at least portions thereof.

Another example of such a first criterion is the relative homogeneity of the vicinity of a pixel point that is being viewed. The assumption is made that homogeneous image areas do not contain any recognizable objects. The vicinity of a pixel point is defined by a given raster pattern in the image. Such vicinity is made up of a number of pixels. In order to ascertain the homogeneity of a pixel point vicinity that is being viewed, pixel points that have similar characteristics are considered. For example, the pixels of an area or vicinity may have similar color values. If these characteristics of the area correspond to the given homogeneity criteria, then the respective pixel

point is classified as not relevant for the performance of the following method steps and is thus not processed. If the given criterion, such as homogeneity is not satisfied, in other words, if a nonhomogeneous vicinity is recognized around a particular pixel point, the respective pixel point is classified as being relevant for the further processing.

10

DEUTSCHE
PATENT-
OBER-
BEAMTE

15

20

Fig. 2 shows an example of an input image 1 that was taken by an infrared camera in a low flying aircraft. The input image 1 shows a scene including a terrain with vehicles. These vehicles are automatically recognizable by means of the method according to the invention. In the first step each pixel point of the image is checked whether the particular pixel point has a vicinity with similar color values. Thus, for example homogeneous image regions or areas are recognized if they are made up of pixel points having rather similar color values. On the other hand, the vehicles show locally distinctly different color values and the respective pixel points do not satisfy the predetermined homogeneity criteria. Thus, all pixel points which, for example belong to fields are classified in the rough classification step 10 as being irrelevant for the further processing while all pixel points that relate, for example to the vehicles are classified for further processing in the following method steps.

25

Referring further to Fig. 2, the right-hand side of Fig. 2 shows the reduced image 11 resulting from the rough classification.

The left-hand side of Fig. 2 showing the input image 1 is the result of an image acquisition by an infrared sensor or camera showing a terrain with a road and dark spots on the road in the lower one half of the left side of Fig. 2. These dark spots represent vehicles. As a result of the rough classification the irrelevant image points have been identified and the right-hand part of Fig. 2 shows these irrelevant pixel points as white areas. These white areas are ignored in the following method steps 20, 30, 40 and 50 in which steps exclusively only the image areas are further processed which are shown as dark pixel points.

The rough classification step 10 shown in Fig. 1 reduces the areas of the input image 1 that need to be processed to form a reduced image 11. The following method steps are concentrated exclusively to the processing of image pixel points remaining in the reduced image 11 following the rough classification step 10. Thus, in the following method steps 20, 30, 40 and 50 the relevant pixel points are exclusively processed and pixel points that were classified as not relevant in step 10 are completely ignored.

The next step 20 involves a dissecting or filtering of the reduced image 11 to form signal presentations by way of preferably multiple filtering substeps in accordance with known section criteria to form several filter images 21, 22 and 23, each of which contains the scene that remained in the reduced image 11 as far as that scene is relevant for the object

recognition. In other words, the filter images 21, 22 and 23 correspond to each other as far as the shown image content is concerned. Generally, the reduced image 11 is divided at least into two corresponding filter images. However, in accordance with the embodiment of Fig. 1, three corresponding filter images 21, 22 and 23 are formed in the filtering step 20. This filtering is preferably performed in a two-dimensional fashion. However, the filtering can also be performed in a one-dimensional fashion, whereby the filtering takes place along the rows or columns of the reduced image 11. The filter images 21, 22, 23 correspond preferably to complementary informations taken from the reduced images 11. These complementary informations contained in the filter images 21, 22, 23 taken together, permit forming or obtaining the complete information gathered from the reduced image 11. In other words, all image components that remained in the image 11 after reduction are subjected to the filtering step 20.

For example, a reduced image 11 could be divided in the filtering step 20 into a first corresponding filter image 21 containing small image elements, a second corresponding filter image 22 with somewhat large image elements, and a third corresponding filter image 23 with remaining still larger image elements. For example, the small image elements could be compact cars, the second elements could be vans, and the third elements could be large trucks. The scene represented in the reduced image 11 is not affected by such filtering. It is retained, which means that

the relevant image components and their correlation relative to one another is retained. Irrelevant image components such as the nature of the terrain are no longer present in the filter images, but the rough scene of the reduced image 11 is fully retained in
5 the filter images 21, 22 and 23. Viewing, for example, the above mentioned image of a road with vehicles on the road and fields next to the road, the filter image 21, for example, could show the object details of the vehicles, such as the tires and the headlights. The filter image 22 could emphasize the central
10 vehicle structural component such as the windshield, the hood, and so forth. The filter image 23 could then emphasize the rough contours of the vehicle. However, in spite of such emphasis of detail each of the filter images 21, 22, 23 keeps containing the entire scene of the road and the vehicles.

15 According to another example embodiment of the invention the filtering step 20 may divide the reduced image 11 in accordance with color areas or in accordance with gray values. Another criterion for the division or filtering of the reduced image 11 could be to form an image pyramid that contains the individual
20 corresponding filter images or representations. Each step of the image pyramid corresponds to another generally lower resolution of the input image 1. Thus, the informations contained in the reduced image 11 are divided in accordance with different resolutions. More specifically, the object or objects contained
25 in the reduced image 11 may be shown in the filter images with different resolutions, whereby each corresponding filter image

shows the respective image information or image object in a different size dimension. For example, passenger cars may be enlarged while trucks are relatively reduced in size.

The next step 30 according to the invention tests each pixel point of each of the filter images 21, 22, 23 with reference to a respective classifier. At least one classifier is provided for each filter image. The total number of classifiers used forms an ensemble or group of classifiers. Such ensemble or group comprises at least two classifiers. According to the invention each classifier operates independently of any other classifier and the classifiers are generally differentiated from one another.

10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95

20
25

The result of the step 30 or the result of the testing of the pixel points by the ensemble of classifiers provides for each filter image 21, 22, 23 a respective group of classification images 31A, 31B, 31C and 32A, 32B, 32C and 33A, 33B, 33C as shown in Fig. 1. In the illustration of Fig. 1 the classification images 31A, 32A and 33A represent the result of the testing with an ensemble of classifiers for a certain object class C1, for example representing a defined vehicle type. The classification images 31B, 32B and 33B represent an object class C2 while the classification images 31C, 32C and 33C show a third object class C3, for example a further different vehicle type. However, the present method is not limited to these three example object classes.

4028/WGF:ar

According to the invention different weighting factors or significance factors are allocated to each pixel point of each of the individual classification images. These weighting factors represent a measure or evaluation of the respective classifier that determines to which object class the pixel point belongs which pixel point is being considered. These weighting factors or evaluation numbers may, for example represent probabilities or a predetermined rank order or hierarchy.

Each classifier operates as follows. Each classifiers ranks the input values, that is each pixel point of the respective pixel images 21, 22, 23 with regard to the pixel point vicinity. More specifically, each classifier allocates to each pixel point of the respective filter image in accordance with a predetermined rule, output values in the form of evaluation numbers or criteria. The output values of the classifier indicate approximately how certain the classifier is with regard to the question: To which class does this pixel point under consideration belong? The following examples of classifiers may be used in accordance to the invention, such as polynomial classifiers, support vector machines, neural networks or the like.

If, for example, neural networks are used as classifiers, Fig. 1 requires an ensemble of three different neural networks. Each of the filter images 21, 22, 23 is allocated to one of these three neural networks. Thus, each neural network processes one

of the filter images. Each neural network allocates to each pixel point of the respective filter image 21, 22, 23 an evaluation measurer or number in accordance with given or predetermined rules for each object class. The evaluation number or measure may, for example be a probability value in the respective classified image 31A, 31B, 31C or 32A, 32B, 32C or 33A, 33B, 33C. Stated differently, the neural network stores the allocated evaluation number or measure and then visualizes that value, for example by a color coding in the classification or 10 classified images 31A, 31B, 31C or 32A, 32B, 32C or 33A, 33B, 33C.

The rule by which a classifier works is preferably obtained from available and analyzed examples of representative input values which have been analyzed prior to the application of the method according to the invention as shown in Fig. 1. This formation of the classifier in accordance with the just mentioned rule generally involves an iterative adaptation of the classifier to the given examples which contain the applicable rule or which define the applicable rule. If neural networks are used this 20 iterative adaption phase is referred to in the respective literature as "training phase" or as "learning of a rule". If other classifiers are employed according to the invention they may, for example, involve a statistical method.

If, for example it is required that different vehicle types must 25 be classified from available input images 1, the classification

step 30 may be based on predetermined classes such as "compact cars", "limousines", and "vans or pick ups". Different features of vehicles are extracted from the reduced images 11 by the filtering step 20. A feature vector or characterizing vector is formed from these features for each pixel point. The components of such a characterizing vector are supplied to the classifiers as input values. The classifier such as a neural network uses these input values as a basis for the classification of each pixel point. More specifically, the classifier allocates, in accordance with a previously determined rule, the above mentioned evaluation number, which may also be referred to as weighting factor or significance number, to each feature vector for each of the classes "compact cars", "limousines", "vans" or "pick ups".

As mentioned, the characterizing or feature vectors required for the training of the neural network classifiers prior to applying the method according to the invention can, for example, be obtained on the basis of available representative example input images. In the forming of these feature or characterizing vectors that are required for the training of the neural network for learning the rules, only those pixel points are used from the filter images 21, 22, 23 which passed the rough classification step 10 of example input images and were recognized or classified as relevant pixel points. The example input images are to be distinguished from the input images 1 that are processed according to the present method.

The input values to each classifier are assembled on the basis of the vicinity of the image or pixel point under consideration in the respective filter image 21, 22 or 23. More specifically, for each pixel point a vicinity around the pixel point is selected. The vicinity includes left, right, upper and lower neighboring pixel points of the pixel point that is being evaluated. The selection of the vicinity including the number of pixel points forming the vicinity can be made dependent on the objects to be recognized or the selection may be fixed. The selection criteria are based on the characteristics or features of the objects to be recognized. For example, large objects require the examination of large vicinities because the recognition of large objects requires generally more features than are necessary for the recognition of smaller objects. The pixel points making up the vicinity are sorted into a feature or characterizing vector in accordance with given rules. Based on this characterizing or feature vector, the respective classifier of the ensemble provides output values for each pixel point and these output values are interpreted with regard to the evaluation number or weighting factor that allocates a relevant pixel point to a specific class.

An example of a rule for the formation of a feature vector from the vicinity of a pixel point will now be described. When forming the characterizing or feature vector, the values representing the vicinity of the pixel point in the respective filter images 21, 22 and 23 are sorted in spiral fashion into a

coefficient vector. A rapid Fourier transformation is then applied to the coefficient vector. Thereafter, the feature vector is formed from the absolute values of the Fourier transformation coefficients.

5 A classification step 30 is performed following the filtering step 20. Step 20 provides signal representations in the form of filter images 21, 22 and 23 which contain features of small, medium sized and large objects, then step 30 is performed in the following sequence. A feature vector is formed for each pixel
10 point of the respective filter image 21, 22, 23. This feature vector is formed from the pixel points forming the vicinity of the pixel point under consideration. For example, a vicinity may be formed by 5×5 pixel points. The pixel point under consideration, which is to be evaluated is in the center of this
15 cluster of pixel points forming the vicinity. This feature vector comprises 25 components in accordance with the selected 5×5 vicinity. The components of the feature vector are assembled from encoded values of the respective filter image 21,
20 22 or 23. A further possible rule for the production of a feature vector based on the 5×5 pixel vicinity provides, for example, that the columns of the 5×5 pixel vicinity are sequentially sorted into the feature vector. Based on the formed
25 feature vector, the classifier that is allocated to the respective filter image provides output values which constitute an evaluation of the pixel point under consideration and the evaluation determines to which object class the respective pixel

point belongs, for example compact cars, limousines, or pick ups or vans. The feature vectors, for example from the filter image 21 comprise, for example vehicle details which are specific to a particular vehicle class, whereby it becomes possible for the classifier to allocate the respective pixel point to the relevant class. Similarly, or analogously, the feature vectors of the filter image 22 have reference to vehicle structures of midsized vehicles and the feature vectors of the filter image 23 have relevance to large vehicle structures. In all instances these vectors form the basis for the decision made by the respective classifiers. The classifier provides for each object class an output value. Such an output value provides, for example information regarding the probability with which a pixel point under consideration can be allocated to a specific object class. This output value is stored in the respective classification image 31A, 32A, 33A, or 31B, 32B, 33B or 31C, 32C, 33C. The output value may be visualized by a color coding of a respective value range. For example, if the output value is made black, the probability is zero percent, a white color for the output value represents, for example a probability of 100%. Assuming that the evaluation numbers are to be expressed in probabilities, it would be possible, for example, that the first classifier which is allocated to the filter image 21 provides, based on the respective feature vector of a certain pixel point the following probability allocations: 13% for the object class compact cars, 17% for the object class limousine, and 70% for the object class pick up or van. These three probabilities are stored in the

classification images 31A, 31B, 31C and visualized. The classification images 32A, 32B, 33C and 33A, 33B, 33C are formed analogously as just described.

The next step 40 in the method according to the invention 5 performs a fusion or merging of the above classification images 31A to 33C for each object class. This merging step combines the obtained individual evaluation numbers of the classifiers for each image or pixel point and for each classification image 31A to 33C to form a global evaluation or decision which is presented 10 in the form of the merged images 41A, 41B and 41C as shown in Fig. 1. In this way one obtains for each image point of the reduced image 11 and for each object class a global evaluation or decision number which represents the combined evaluation of the classifiers forming a group or classifier ensemble.

The individual evaluation numbers of the classifiers of an ensemble are combined in accordance with predetermined known mathematical methods. For example a mean value may be calculated 15 on the basis of the individual evaluation numbers and this mean value may be used as the global evaluation result. If one considers the decision of a classifier as a probability statement, statistical methods may be used in order to obtain the final or global decision. Such statistical methods are, for 20 example the Bayes-Fusion or the Dempster-Shafer-Fusion. For this purpose the output values of the classifiers are approximated in accordance with probabilities and are then merged with the aid 25

of the probability theory and on the basis of known apriori probabilities. The apriori probabilities may, for example, be obtained from context information regarding the type, position and content of the input image 1 to be evaluated. The 5 probability values of the classifiers or the ensemble of classifiers are gathered pixel by pixel for each object class in the merging or fusing step 40. The merging or fusing result of each object class thereby corresponds to a final probability statement for the pixel point under consideration or to be 10 evaluated and with reference to the predetermined object class.

The merging or fusion step 40 for the individual evaluation numbers of the ensemble for any random pixel point of the reduced image 11 may, for example take place in the following sequence. The evaluation numbers of the different classifiers of the corresponding pixel points of the classification images 31A to 33C are combined with an algorithm for each object class. Assuming, for example that the evaluation numbers are probabilities then the classifiers of the ensemble provide for a certain object class and pixel by pixel a respective 20 probability percentage, for example the following probabilities: 87%, 83% and 95%. The merging or fusing result for this pixel point and for this object class could now be determined, for example as the mean value of the three probabilities, thereby obtaining 88.3%. The individual numbers of probabilities in this 25 example have the following significance or meaning: a first classifier of the ensemble which is allocated to the filter image

21 provides for a defined object class and for the pixel point
of the reduced image 11 under consideration the probability value
of 87%. More specifically, the classifier allocates to the pixel
point under consideration and on the basis of the features of the
filter image 21 the probability of 87% that this pixel point
belongs to a defined object class, for example object class A
represented by the cluster A in Fig. 3. The probability values
of the remaining classifiers and pixel points are to be
interpreted correspondingly. The fusion or merged overall or
10 global probability per pixel point means that the ensemble of
image classifiers has allocated to the pixel point under
consideration a probability of, for example 88.3% that it belongs
to a determined object class such as object class A. The merging
or fusing for the other object classes takes place in an
analogous manner.

The last step 50 according to the invention establishes a
decision result image 51. In this step it is determined on the
basis of the merging or fusion images 41A, 41B, 41C for each
pixel point of the reduced image 11 whether that point belongs
20 to an object class and, if so, to which of these object classes
the pixel point belongs. For this allocation in step 50 the
total evaluation numbers of the corresponding pixel points of the
fusion images 41A, 41B, 41C are combined with each other, for
example by forming the maximum of the total evaluation numbers
25 for the different object classes. Then that object class is
selected which has the highest value in the global evaluation

number. The thus determined maximum global evaluation number for one object class is then compared with a given threshold value. If the global evaluation member or value exceeds the threshold value a corresponding class, suitably encoded, is allocated to the respective pixel point in the decision result image 51. The previously selected threshold value determines for each pixel point whether or not that pixel point is part of an object of a determined object class.

A possible formation of the decision result image 51 will now be described. Assuming that three recognizable object classes are used, for example compact cars, limousines, pick ups or vans. Under this assumption a certain pixel point, for example of the object class "vans" has assigned thereto the highest value of the respective global evaluation number. If this highest value is larger than a given threshold value, then the pixel point under consideration is allocated to the object class "vans". If the global evaluation number is below the threshold value, it is assumed that the respective pixel point does not belong to any of the object classes. Zero (0) have been allocated to the pixel points that have been eliminated in the rough classification 10, as shown in Fig. 3. Identification numbers 1, 2, 3 have been allocated to the pixel points that fall into the three example object classes. These pixel points are part of the decision result image 51. For example, "1" refers to the object class of compact cars, while "2" refers to the class of limousines and "3"

refers to vans and pick ups as described above and shown in Fig. 3. The letters A, B, C designate respective pixel clusters.

Referring further to Fig. 3, the decision result image 51 illustrates recognized objects in the three object classes represented by respective three pixel clusters A, B and C. A pixel cluster is generally formed by a larger number of neighboring pixels with identical characteristics, for example identical color values. In the decision result image 51 the pixel coding numbers 0, 1, 2 and 3 are given merely as an example. An object class is characterized in that the neighboring pixel points are predominantly allocated to a certain object class so that these pixel points predominantly have the same encoding. For example, clusters A and B in Fig. 3 are uniformly formed by code numbers 1 and 2, respectively. However, cluster C has within a plurality of pixels that are encoded with number 3, one pixel point encoded with number 1. These clusters correspond to defined vehicle types as described. The size of the respective cluster corresponds approximately to the dimensions of the vehicle or vehicles in the initial input image 1. The fact that the cluster C in Fig. 3 includes one pixel that does not belong into the respective object class does not make the cluster C ambiguous.

The decision result image 51 may be used for further processing in the object plane. However, the method according to the invention relates only to the image pixel plane. In the

following an example is provided how the transition from the iconic or image pixel plane to the symbolic object plane can be performed. In the symbolic object plane the image content is described by symbols. The method steps for such transition from the iconic image pixel plane to the symbolic object plane are well known in the art. Based on, for example the cluster size that is the number of pixels having the same encoding in a common cluster, it is possible to decide in the decision image 51 whether or not any particular cluster represents a real object falling into a predetermined object class in the input image 1, on the symbolic object plane. If, for example, the pixel points form a cluster in the decision result image 51, the size of which exceeds a certain value, for example more than 100 pixels, the cluster is considered to be equal to a recognized object of the respective object class. The positions of the recognized objects in the input image 1 may, for example be approximated from the positions of cluster concentration in the decision result image 51.

The reduced image 11, the filter images 21, 22, 23, the classification images 31A, 32A, 33A or 31B, 32B, 33B or 31C, 32C, 33C, the fusion or merging images 41A, 41B, and 41C, as well as the decision result image 51 need not be imaged in a visual sense. Rather, these images may be formulated by means of mathematical methods, for example by means of matrices.

Fig. 4 shows the method according to the invention as performed with reference to only one relevant object class for which a search is made. Fig. 4 illustrates the simplest application of the present invention. However, the present method may be used
5 for recognizing objects in a plurality of object classes.

Although the invention has been described with reference to specific example embodiments, it will be appreciated that it is intended to cover all modifications and equivalents within the scope of the appended claims. It should also be understood that
10 the present disclosure includes all possible combinations of any individual features recited in any of the appended claims.

SEARCHED
INDEXED
SERIALIZED
FILED